1631

ENTERED

1600

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/056,019B

DATE: 01/06/2003
TIME: 13:27:17

#35

Input Set : A:\5853-2 Sequence Listing .txt
Output Set: N:\CRF4\01062003\I056019B.raw

RECEIVED
JAN 15 2003
TECH CENTER 1600/2900

<110> APPLICANT: Tuomanen, Elaine I
Wizemann, Theresa M.
Masure, H. R.
Johnson, Leslie S.
Koenig, Scott
<120> TITLE OF INVENTION: POLYPEPTIDE COMPRISING THE AMINO ACID OF AN N-TERMINAL CHOLINE BINDING PROTEIN A TRUNCATE, VACCINE DERIVED THEREFROM AND USES THEREOF
<130> FILE REFERENCE: 5853-2
<140> CURRENT APPLICATION NUMBER: 09/056,019B
<141> CURRENT FILING DATE: 1998-04-07
<160> NUMBER OF SEQ ID NOS: 40
<170> SOFTWARE: PatentIn Ver. 2.0
<210> SEQ ID NO: 1
<211> LENGTH: 406
<212> TYPE: PRT
<213> ORGANISM: Streptococcus pneumoniae
<400> SEQUENCE: 1

```
Glu Asn Glu Gly Ala Thr Gln Val Pro Thr Ser Ser Asn Arg Ala Asn
 1               5                  10                  15
Glu Ser Gln Ala Glu Gln Gly Glu Gln Pro Lys Lys Leu Asp Ser Glu
            20                  25                  30
Arg Asp Lys Ala Arg Lys Glu Val Glu Glu Tyr Val Lys Lys Ile Val
        35                  40                  45
Gly Glu Ser Tyr Ala Lys Ser Thr Lys Lys Arg His Thr Ile Thr Val
    50                  55                  60
Ala Leu Val Asn Glu Leu Asn Asn Ile Lys Asn Glu Tyr Leu Asn Lys
65                  70                  75                  80
Ile Val Glu Ser Thr Ser Glu Ser Gln Leu Gln Ile Leu Met Met Glu
                85                  90                  95
Ser Arg Ser Lys Val Asp Glu Ala Val Ser Lys Phe Glu Lys Asp Ser
            100                 105                 110
Ser Ser Ser Ser Ser Ser Asp Ser Ser Thr Lys Pro Glu Ala Ser Asp
        115                 120                 125
Thr Ala Lys Pro Asn Lys Pro Thr Glu Pro Gly Glu Lys Val Ala Glu
    130                 135                 140
Ala Lys Lys Lys Val Glu Glu Ala Glu Lys Lys Ala Lys Asp Gln Lys
145                 150                 155                 160
Glu Glu Asp Arg Arg Asn Tyr Pro Thr Ile Thr Tyr Lys Thr Leu Glu
                165                 170                 175
Leu Glu Ile Ala Glu Ser Asp Val Glu Val Lys Lys Ala Glu Leu Glu
            180                 185                 190
Leu Val Lys Val Lys Ala Asn Glu Pro Arg Asp Glu Gln Lys Ile Lys
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/056,019B

DATE: 01/06/2003
TIME: 13:27:17

Input Set : A:\5853-2 Sequence Listing .txt
Output Set: N:\CRF4\01062003\I056019B.raw

```
65          195                 200                 205
67 Gln Ala Glu Ala Glu Val Glu Ser Lys Gln Ala Glu Ala Thr Arg Leu
68      210                 215                 220
70 Lys Lys Ile Lys Thr Asp Arg Glu Glu Ala Glu Glu Glu Ala Lys Arg
71 225                 230                 235                 240
73 Arg Ala Asp Ala Lys Glu Gln Gly Lys Pro Lys Gly Arg Ala Lys Arg
74                  245                 250                 255
76 Gly Val Pro Gly Glu Leu Ala Thr Pro Asp Lys Lys Glu Asn Asp Ala
77              260                 265                 270
79 Lys Ser Ser Asp Ser Ser Val Gly Glu Glu Thr Leu Pro Ser Pro Ser
80          275                 280                 285
82 Leu Lys Pro Glu Lys Lys Val Ala Glu Ala Glu Lys Lys Val Glu Glu
83      290                 295                 300
85 Ala Lys Lys Lys Ala Glu Asp Gln Lys Glu Glu Asp Arg Arg Asn Tyr
86 305                 310                 315                 320
88 Pro Thr Asn Thr Tyr Lys Thr Leu Glu Leu Glu Ile Ala Glu Ser Asp
89                  325                 330                 335
91 Val Glu Val Lys Lys Ala Glu Leu Glu Leu Val Lys Glu Glu Ala Lys
92              340                 345                 350
94 Glu Pro Arg Asn Glu Glu Lys Val Lys Gln Ala Lys Ala Glu Val Glu
95          355                 360                 365
97 Ser Lys Lys Ala Glu Ala Thr Arg Leu Glu Lys Ile Lys Thr Asp Arg
98      370                 375                 380
100 Lys Lys Ala Glu Glu Glu Ala Lys Arg Lys Ala Ala Glu Glu Asp Lys
101 385                 390                 395                 400
103 Val Lys Glu Lys Pro Ala
104                 405
107 <210> SEQ ID NO: 2
108 <211> LENGTH: 655
109 <212> TYPE: PRT
110 <213> ORGANISM: Streptococcus pneumoniae
112 <400> SEQUENCE: 2
113 Glu Asn Glu Gly Ala Thr Gln Val Pro Thr Ser Ser Asn Arg Ala Asn
114   1               5                  10                  15
116 Glu Ser Gln Ala Glu Gln Gly Glu Gln Pro Lys Lys Leu Asp Ser Glu
117             20                  25                  30
119 Arg Asp Lys Ala Arg Lys Glu Val Glu Glu Tyr Val Lys Lys Ile Val
120         35                  40                  45
122 Gly Glu Ser Tyr Ala Lys Ser Thr Lys Lys Arg His Thr Ile Thr Val
123     50                  55                  60
125 Ala Leu Val Asn Glu Leu Asn Asn Ile Lys Asn Glu Tyr Leu Asn Lys
126 65                  70                  75                  80
128 Ile Val Glu Ser Thr Ser Glu Ser Gln Leu Gln Ile Leu Met Met Glu
129                 85                  90                  95
131 Ser Arg Ser Lys Val Asp Glu Ala Val Ser Lys Phe Glu Lys Asp Ser
132             100                 105                 110
134 Ser Ser Ser Ser Ser Ser Asp Ser Ser Thr Lys Pro Glu Ala Ser Asp
135         115                 120                 125
137 Thr Ala Lys Pro Asn Lys Pro Thr Glu Pro Gly Glu Lys Val Ala Glu
```

RAW SEQUENCE LISTING DATE: 01/06/2003
PATENT APPLICATION: **US/09/056,019B** TIME: 13:27:17

Input Set : **A:\5853-2 Sequence Listing .txt**
Output Set: **N:\CRF4\01062003\I056019B.raw**

```
138     130                 135                 140
140 Ala Lys Lys Lys Val Glu Glu Ala Glu Lys Lys Ala Lys Asp Gln Lys
141 145                 150                 155                 160
143 Glu Glu Asp Arg Arg Asn Tyr Pro Thr Ile Thr Tyr Lys Thr Leu Glu
144                 165                 170                 175
146 Leu Glu Ile Ala Glu Ser Asp Val Glu Val Lys Lys Ala Glu Leu Glu
147             180                 185                 190
149 Leu Val Lys Val Lys Ala Asn Glu Pro Arg Asp Glu Gln Lys Ile Lys
150         195                 200                 205
152 Gln Ala Glu Ala Glu Val Glu Ser Lys Gln Ala Glu Ala Thr Arg Leu
153     210                 215                 220
155 Lys Lys Ile Lys Thr Asp Arg Glu Glu Ala Glu Glu Glu Ala Lys Arg
156 225                 230                 235                 240
158 Arg Ala Asp Ala Lys Glu Gln Gly Lys Pro Lys Gly Arg Ala Lys Arg
159                 245                 250                 255
161 Gly Val Pro Gly Glu Leu Ala Thr Pro Asp Lys Lys Glu Asn Asp Ala
162             260                 265                 270
164 Lys Ser Ser Asp Ser Ser Val Gly Glu Glu Thr Leu Pro Ser Pro Ser
165         275                 280                 285
167 Leu Lys Pro Glu Lys Lys Val Ala Glu Ala Glu Lys Lys Val Glu Glu
168     290                 295                 300
170 Ala Lys Lys Lys Ala Glu Asp Gln Lys Glu Glu Asp Arg Arg Asn Tyr
171 305                 310                 315                 320
173 Pro Thr Asn Thr Tyr Lys Thr Leu Glu Leu Glu Ile Ala Glu Ser Asp
174                 325                 330                 335
176 Val Glu Val Lys Lys Ala Glu Leu Glu Leu Val Lys Glu Glu Ala Lys
177             340                 345                 350
179 Glu Pro Arg Asn Glu Glu Lys Val Lys Gln Ala Lys Ala Glu Val Glu
180         355                 360                 365
182 Ser Lys Lys Ala Glu Ala Thr Arg Leu Glu Lys Ile Lys Thr Asp Arg
183     370                 375                 380
185 Lys Lys Ala Glu Glu Glu Ala Lys Arg Lys Ala Ala Glu Glu Asp Lys
186 385                 390                 395                 400
188 Val Lys Glu Lys Pro Ala Glu Gln Pro Gln Pro Ala Pro Ala Pro Lys
189                 405                 410                 415
191 Ala Glu Lys Pro Ala Pro Ala Pro Lys Pro Glu Asn Pro Ala Glu Gln
192             420                 425                 430
194 Pro Lys Ala Glu Lys Pro Ala Asp Gln Gln Ala Glu Glu Asp Tyr Ala
195         435                 440                 445
197 Arg Arg Ser Glu Glu Glu Tyr Asn Arg Leu Thr Gln Gln Gln Pro Pro
198     450                 455                 460
200 Lys Thr Glu Lys Pro Ala Gln Pro Ser Thr Pro Lys Thr Gly Trp Lys
201 465                 470                 475                 480
203 Gln Glu Asn Gly Met Trp Tyr Phe Tyr Asn Thr Asp Gly Ser Met Ala
204                 485                 490                 495
206 Thr Gly Trp Leu Gln Asn Asn Gly Ser Trp Tyr Tyr Leu Asn Ser Asn
207             500                 505                 510
209 Gly Ala Met Ala Thr Gly Trp Leu Gln Asn Asn Gly Ser Trp Tyr Tyr
210         515                 520                 525
```

```
212 Leu Asn Ala Asn Gly Ser Met Ala Thr Gly Trp Leu Gln Asn Asn Gly
213     530                 535                 540
215 Ser Trp Tyr Tyr Leu Asn Ala Asn Gly Ser Met Ala Thr Gly Trp Leu
216 545                 550                 555                 560
218 Gln Tyr Asn Gly Ser Trp Tyr Tyr Leu Asn Ala Asn Gly Ser Met Ala
219                 565                 570                 575
221 Thr Gly Trp Leu Gln Tyr Asn Gly Ser Trp Tyr Tyr Leu Asn Ala Asn
222             580                 585                 590
224 Gly Asp Met Ala Thr Gly Trp Val Lys Asp Gly Asp Thr Trp Tyr Tyr
225         595                 600                 605
227 Leu Glu Ala Ser Gly Ala Met Lys Ala Ser Gln Trp Phe Lys Val Ser
228     610                 615                 620
230 Asp Lys Trp Tyr Tyr Val Asn Gly Ser Gly Ala Leu Ala Val Asn Thr
231 625                 630                 635                 640
233 Thr Val Asp Gly Tyr Gly Val Asn Ala Asn Gly Glu Trp Val Asn
234                 645                 650                 655
237 <210> SEQ ID NO: 3
238 <211> LENGTH: 284
239 <212> TYPE: PRT
240 <213> ORGANISM: Streptococcus pneumoniae
242 <400> SEQUENCE: 3
243 Glu Asn Glu Gly Ala Thr Gln Val Pro Thr Ser Ser Asn Arg Ala Asn
244   1               5                  10                  15
246 Glu Ser Gln Ala Glu Gln Gly Glu Gln Pro Lys Lys Leu Asp Ser Glu
247             20                  25                  30
249 Arg Asp Lys Ala Arg Lys Glu Val Glu Glu Tyr Val Lys Lys Ile Val
250         35                  40                  45
252 Gly Glu Ser Tyr Ala Lys Ser Thr Lys Lys Arg His Thr Ile Thr Val
253     50                  55                  60
255 Ala Leu Val Asn Glu Leu Asn Asn Ile Lys Asn Glu Tyr Leu Asn Lys
256 65                  70                  75                  80
258 Ile Val Glu Ser Thr Ser Glu Ser Gln Leu Gln Ile Leu Met Met Glu
259                 85                  90                  95
261 Ser Arg Ser Lys Val Asp Glu Ala Val Ser Lys Phe Glu Lys Asp Ser
262             100                 105                 110
264 Ser Ser Ser Ser Ser Ser Asp Ser Ser Thr Lys Pro Glu Ala Ser Asp
265         115                 120                 125
267 Thr Ala Lys Pro Asn Lys Pro Thr Glu Pro Gly Glu Lys Val Ala Glu
268     130                 135                 140
270 Ala Lys Lys Lys Val Glu Glu Ala Glu Lys Lys Ala Lys Asp Gln Lys
271 145                 150                 155                 160
273 Glu Glu Asp Arg Arg Asn Tyr Pro Thr Ile Thr Tyr Lys Thr Leu Glu
274                 165                 170                 175
276 Leu Glu Ile Ala Glu Ser Asp Val Glu Val Lys Lys Ala Glu Leu Glu
277             180                 185                 190
279 Leu Val Lys Val Lys Ala Asn Glu Pro Arg Asp Glu Gln Lys Ile Lys
280         195                 200                 205
282 Gln Ala Glu Ala Glu Val Glu Ser Lys Gln Ala Glu Ala Thr Arg Leu
283     210                 215                 220
```

RAW SEQUENCE LISTING DATE: 01/06/2003
PATENT APPLICATION: **US/09/056,019B** TIME: 13:27:17

Input Set : **A:\5853-2 Sequence Listing .txt**
Output Set: **N:\CRF4\01062003\I056019B.raw**

```
285 Lys Lys Ile Lys Thr Asp Arg Glu Glu Ala Glu Glu Glu Ala Lys Arg
286 225                 230                 235                 240
288 Arg Ala Asp Ala Lys Glu Gln Gly Lys Pro Lys Gly Arg Ala Lys Arg
289                 245                 250                 255
291 Gly Val Pro Gly Glu Leu Ala Thr Pro Asp Lys Lys Glu Asn Asp Ala
292             260                 265                 270
294 Lys Ser Ser Asp Ser Ser Val Gly Glu Glu Thr Leu
295         275                 280
298 <210> SEQ ID NO: 4
299 <211> LENGTH: 106
300 <212> TYPE: PRT
301 <213> ORGANISM: Streptococcus pneumoniae
303 <400> SEQUENCE: 4
304 Lys Pro Glu Lys Lys Val Ala Glu Ala Glu Lys Lys Val Glu Glu Ala
305   1               5                  10                  15
307 Lys Lys Lys Ala Glu Asp Gln Lys Glu Glu Asp Arg Arg Asn Tyr Pro
308             20                  25                  30
310 Thr Asn Thr Tyr Lys Thr Leu Glu Leu Glu Ile Ala Glu Ser Asp Val
311          35                  40                  45
313 Glu Val Lys Lys Ala Glu Leu Glu Leu Val Lys Glu Glu Ala Lys Glu
314      50                  55                  60
316 Pro Arg Asn Glu Glu Lys Val Lys Gln Ala Lys Ala Glu Val Glu Ser
317  65                  70                  75                  80
319 Lys Lys Ala Glu Ala Thr Arg Leu Glu Lys Ile Lys Thr Asp Arg Lys
320                 85                  90                  95
322 Lys Ala Glu Glu Glu Ala Lys Arg Lys Ala
323             100                 105
326 <210> SEQ ID NO: 5
327 <211> LENGTH: 109
328 <212> TYPE: PRT
329 <213> ORGANISM: Streptococcus pneumoniae
331 <400> SEQUENCE: 5
332 Thr Glu Pro Gly Glu Lys Val Ala Glu Ala Lys Lys Lys Val Glu Glu
333   1               5                  10                  15
335 Ala Glu Lys Lys Ala Lys Asp Gln Lys Glu Glu Asp Arg Arg Asn Tyr
336             20                  25                  30
338 Pro Thr Ile Thr Tyr Lys Thr Leu Glu Leu Glu Ile Ala Glu Ser Asp
339          35                  40                  45
341 Val Glu Val Lys Lys Ala Glu Leu Glu Leu Val Lys Val Lys Ala Asn
342      50                  55                  60
344 Glu Pro Arg Asp Glu Gln Lys Ile Lys Gln Ala Glu Ala Glu Val Glu
345  65                  70                  75                  80
347 Ser Lys Gln Ala Glu Ala Thr Arg Leu Lys Lys Ile Lys Thr Asp Arg
348                 85                  90                  95
350 Glu Glu Ala Glu Glu Glu Ala Lys Arg Arg Ala Asp Ala
351             100                 105
354 <210> SEQ ID NO: 6
355 <211> LENGTH: 4
356 <212> TYPE: PRT
```

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 01/06/2003
PATENT APPLICATION: **US/09/056,019B** TIME: 13:27:18

Input Set : **A:\5853-2 Sequence Listing .txt**
Output Set: **N:\CRF4\01062003\I056019B.raw**

Please Note:
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:6; Xaa Pos. 2,3
Seq#:27; Xaa Pos. 1
Seq#:28; Xaa Pos. 243

VERIFICATION SUMMARY DATE: 01/06/2003
PATENT APPLICATION: **US/09/056,019B** TIME: 13:27:18

Input Set : **A:\5853-2 Sequence Listing .txt**
Output Set: **N:\CRF4\01062003\I056019B.raw**

L:366 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6 after pos.:0
L:1099 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:27 after pos.:0
L:1159 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:28 after pos.:240